

Die Aussiedler-Stichprobenziehung

Salentin, Kurt

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Salentin, K. (2007). Die Aussiedler-Stichprobenziehung. *Methoden, Daten, Analysen (mda)*, 1(1), 25-44. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-126439>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Die Aussiedler-Stichprobenziehung

Sampling *Aussiedler*

Kurt Salentin

Zusammenfassung

Deutschstämmige Aussiedler bilden die größte Zuwanderergruppe in der Bundesrepublik. Obwohl Integrationsprobleme bekannt sind, fehlt es an amtstatistischen und Erhebungsdaten, weil eindeutige Identifikationsmerkmale bislang fehlen. Vor- und Nachteile bisheriger Stichprobentechniken für Primärerhebungen (Random Route, Zugangslisten in Übergangsheimen, Schneeballstichproben, onomastische Verfahren) werden erörtert. Der Beitrag stellt dann zum einen ein duales Verfahren vor, das Aussiedler in Einwohnermeldedateien und anderen Registern (a) an spezifischen Doppelstaatsangehörigkeiten und (b) an Geburtsorten erkennt. Er schildert Details des Verfahrens. Anhand einer Anwendung auf Geburtsorte im Bielefelder Melderegister und verschiedener statistischer Quellen wird die Ausschöpfung und der erwartbare Anteil falsch positiver Erkennungen (Beifang) abgeschätzt. Gegenwärtig ist eine Ausschöpfung von ca. 95% der Zielgruppe zu erwarten. Je nach Herkunftsland muss mit Beifang im einstelligen Prozentbereich gerechnet werden, für Polen ist historisch bedingt mit einem erhöhten Anteil zu rechnen. Auf die technisch aufwendige Identifikation des Geburtsortes sollte nicht verzichtet werden, weil ca. die Hälfte der Aussiedler nur die deutsche Staatsangehörigkeit besitzt. Zum anderen wird die Nutzung aussiedlerspezifischer Vor- und Zunamen zur Stichprobenziehung erörtert. In einer Melderegisterstichprobe nach Geburtsort überwiegen deutsche Namen. Allerdings existieren sowohl deutsche Vor- als auch Nachnamen, die Aussiedler relativ gut von autochthonen Deutschen differenzieren. Bis auf ein späteres Einreisejahr und gewisse Abweichungen in Netzwerk- und Kontaktmustern unterscheiden sich die Träger herkunftslandtypischer Namen nicht von den Trägern deutscher Namen.

Abstract

Aussiedler (resettlers of German descent) are the largest immigrant group in contemporary Germany. Though economic and social integration difficulties on their part are known, pertinent survey and official statistical data are still lacking due to the absence of clear-cut criteria discerning them from the autochthonous population. The article discusses advantages and disadvantages of *Aussiedler* sampling techniques hitherto used (random route, address lists kept in transit camps, snowball sampling, name-based techniques). An alternative method based on (a) dual citizenship patterns and (b) specific birthplaces in the population register is then introduced and discussed. Estimates of target population exhaustion and the rate of false identification are calculated from a pilot implementation in the city of Bielefeld. Currently, an exhaustion of 95% may be expected with a rate of false identification of below 10%, varying between former countries of residence. Though cumbersome in technical terms, the birthplace criterion is superior to citizenship-based identification, as about half the *Aussiedler* population cannot be discerned via combinations of German and second citizenships. Further, sampling based on names typical to the group is tested. Though *Aussiedler* names predominantly sound German, a range of names exists that distinguish *Aussiedler* fairly well from autochthonous Germans. Apart from particularities concerning the period of immigration and network and social contact patterns, bearers of German names vs. names typical for countries of former residence do not display significantly different social characteristics.

1 Zielgruppe und bisherige Stichprobenverfahren¹

Der vorliegende Beitrag soll Möglichkeiten und Grenzen der verfügbaren sozialwissenschaftlichen Stichprobenverfahren für deutschstämmige Aussiedler in der Bundesrepublik aufzeigen. Drei Umstände gebieten eine solche Erörterung. 1. Aussiedler bilden hier die größte Einwanderergruppe. Zwischen dem 1.1.1950 und dem 31.12.2004 sind nach Angaben des Bundesverwaltungsamts 4.446.360 Aussiedler einschließlich ihrer Angehörigen ins Bundesgebiet eingereist; seit 1989 sind es 2.823.724 Personen.² Alleine der Umfang der 2.298.938 aus der ehemaligen UdSSR Stammen dieses Zeitraums erreicht beinahe die Stärke der ca. 2.430.000 Türkischstämmigen, die am 31.12.2004 im Bundesgebiet als Ausländer registriert sind (1.764.318 Personen) oder als Eingebürgerte leben (ca. 666.000 Personen).³ 2. In den Sozial- und Bildungswissenschaften herrscht weitgehend Einvernehmen darüber, dass die Platzierung vieler Aussiedler, insbesondere der Spätaussiedler seit 1993, auf den sozialen und ökonomischen Statushierarchien aufgrund sprachlicher und qualifikatorischer Defizite sowie anderer migrantentypischer Handicaps prekär ist (z. B. Seifert 2001). 3. Gleichzeitig bestehen in Wissenschaft und Politik erhebliche Kenntnisdefizite, da Aussiedler als Deutsche in der amtlichen Statistik, etwa der Bildungs-, Sozialhilfe- und mit Ausnahmen auch der Arbeitsmarktstatistik, nicht gesondert ausgewiesen werden. In der quantitativen Sozialforschung gelten Aussiedler als schwer erreichbar, weil Auswahlverfahren, die sich nicht an klaren Kriterien wie der Staatsangehörigkeit orientieren können, keine repräsentativen Aussagen erwarten lassen.

Als Grundgesamtheit sollen hier alle nach dem Bundesvertriebenengesetz als Aussiedler oder Spätaussiedler anerkannten Personen gelten.⁴ Wie immer in Fragen der Migrations- und Minderheitenforschung hängt es von der Fragestellung ab, ob auch die Nachkommen einbezogen werden sollen, d.h. ob eine über die juristische Kategorie hinausgehende Definition sinnvoll ist.

Generell besitzen die Kommunalverwaltungen bestmögliche Ausgangsbedingungen zur Identifikation des Personenkreises mit Migrationshintergrund, da ihnen mehr Informationen zur Verfügung stehen als die Menge der durch das Melderecht zur Übermittlung an Dritte freigegebenen Variablen. Sie sind daher u. a. in der Lage, Aussiedler zu erkennen sowie allgemein aus dem Haushaltskontext einen Migrationshintergrund auch bei solchen Personen zu erschließen, die keine direkten Migrationsindikatoren wie ausländische Geburtsorte oder Staatsangehörigkeiten

1 Der Autor dankt dem Amt für Bürgerberatung in Verbindung mit dem Amt für Stadtplanung, Statistik und Wahlen der Stadt Bielefeld für die Bereitstellung dreier anonymisierter Melderegisterauszüge.

2 Bundesverwaltungsamt, www.bva.bund.de/imperia/md/content/abteilungen/abteilungiii/iist/28.pdf, S. 2. (13.12.2005),
3 Quellen: Statistische Jahrbücher seit 1970, diverse schriftliche Mitteilungen des Statistischen Bundesamts.

4 Nach § 2 (3) in Verbindung mit § 1 (1) des Gesetzes unterscheiden sich Aussiedler von anderen Vertriebenen dadurch, dass sie die deutschen Siedlungsgebiete „nach Abschluß der allgemeinen Vertriebungsmaßnahmen“ verlassen haben. Die administrative Unterscheidung zwischen Aussiedlern und Spätaussiedlern beruht im wesentlichen auf der Ausreise aus der ehemaligen Sowjetunion und dem Baltikum vor dem 31.12.1992. Nichtdeutsche Ehegatten und Abkömmlinge von Aussiedlern werden nach § 4 (3) mit der Einreise in die Bundesrepublik Spätaussiedler. Da die Verreibungen in der Mitte der 1950er Jahre endeten, stellt sich die Frage nach einer Differenzierung von Aussiedlern und anderen Vertriebenen nur bei den heute über 50Jährigen.

besitzen (aus juristischen Gründen nur bei Minderjährigen). Einige deutsche Städte haben begonnen, durch manuelle Einzelfallprüfung den Zuwanderungshintergrund ihrer Einwohner vollständig zu bestimmen (Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration 2005: 551; ergänzt um mündliche Auskünfte eines Vertreters des Deutschen Städtetages vom 16.1.2006). Doch leider wird diese sehr präzise Informationsquelle auf absehbare Zeit nicht für wissenschaftliche Zwecke nutzbar sein. Einerseits sind es meist nicht die Einwohnermeldeämter, sondern die Planungs-, Statistik- und Stadtforschungsämter, die diese wertvolle Pionierarbeit leisten. Im Gegensatz zum Meldewesen verfügt dieser Funktionsbereich über keinerlei Rechtsgrundlage einer Weitergabe von Individualdaten, so dass Forscher allenfalls kleinräumige Aggregatdaten werden nutzen können. (Bereits die juristische Absicherung der Nutzung von Meldedaten durch Planungsämter ist nicht unkompliziert.) Andererseits ist derzeit unklar, ob man sich auf ein einheitliches Konstruktionsverfahren für den Migrationshintergrund mit vergleichbaren Ergebnissen verständigen und inwiefern es angesichts des enormen Aufwands zu einer flächendeckenden Anwendung des Verfahrens kommen wird. Bis auf weiteres müssen Sozialwissenschaftler deshalb andere Wege einschlagen.

Bisher sind im Rahmen quantitativer Studien mehrere Zugangsstrategien für Aussiedler verfolgt worden (auf qualitative Studien kann hier nicht eingegangen werden).

1. Das nach zwei Wellen (1991, 1994) nicht fortgesetzte Aussiedlerpanel des Bundesinstituts für Bevölkerungsforschung (Mamney/Schiener 1996: 146ff., 1998: 11ff.) stellte in nach regionaler Schichtung ausgewählten Übergangsheimen eine Stichprobe von Neuzuwanderern zusammen. In ähnlicher Weise verfuhr Swiaczny (2000) in einer auf Mannheim beschränkten Lokalstudie. Mamney/Schiener diskutierten und verworfen Alternativen: Ein zentrales Aussiedlerregister existiert nicht. In den Einwohnerregistern der kommunalen Meldebehörden ist kein personenbezogenes Merkmal „Aussiedler“ vorhanden. Eine Adressliste des mit dem Lastenausgleichsverfahren befassten Bundesausgleichsamts konnte aus Datenschutzgründen nicht genutzt werden. Das von Befragungsinstituten sonst bevorzugte Random-Route-Verfahren kam (und kommt auch heute) wegen des geringen Anteils von Aussiedlerhaushalten nicht in Frage. Man sah keine Möglichkeit, anstelle einer Zugangsstichprobe eine repräsentative Bestandsstichprobe einschließlich früherer Zugangskohorten zu erstellen. Weil der Erstkontakt bei einer Stichprobe aus Personen ohne eigene Wohnung unternommen wurde, stellte sich enorme Panelmortalität (über 50% von Welle 1 zu Welle 2) ein.
2. Die Aussiedlerhaushalte in der Zuwandererstichprobe des Sozio-oekonomischen Panels (SOEP, s. Schupp/Wagner 1995; Haisken-DeNew/Frick 2005) wurden vom durchführenden Institut zuvor bei anderweitiger Feldarbeit registriert, so dass kein erhöhter Random-Route-Aufwand entstand. Forschern ohne ständige Feldtätigkeit in der Wohnbevölkerung steht dieses Verfahren nicht zur Verfügung. Bei Aussiedler-Analysen bereitet das SOEP zudem unvermeidliche Probleme. Reichen die Daten für manche Forschungszwecke aus, sind sie als Grundlage für Sekundäranalysen doch immer mit inhaltlichen Einschränkungen verbunden. Daneben erzeugt die überwiegend statische Stichprobenanlage des Panels⁵ für eine sich

5 Nur Ausgründungen aus sowie individuelle Zugänge zu bestehenden Haushalten erweitern die Stichprobe.

durch kontinuierliche Zuwanderung dynamisch verändernde Gruppe Abbildungsdefizite eigener Art. So sind etwa die letzten Aussiedlerkohorten mit auch für diese Herkunftsgruppe vergleichsweise schlechten Deutschkenntnissen, unzureichender beruflicher Qualifikation und erhöhtem Armutsrisiko (Engels 2000) im SOEP zwangsläufig kaum vertreten.

3. Auf das manchmal als Notlösung gewählte Schneeballverfahren soll hier wegen dessen bekannter Unsicherheiten (Gabler 1992) nicht ausführlich eingegangen werden. Für die SOEP-Zuwandererstichprobe wurde parallel zu dem gerade beschriebenen Verfahren mit dem Schneeballsystem experimentiert. Schupp/Wagner (1995) beschreiben die Verzerrungen, die aufgrund ungleicher Netzwerkstrukturen auftraten und aufgrund derer dieses Verfahren nicht weiter eingesetzt wurde.
4. Für Untersuchungen mit eigenständiger Stichprobenziehung versprechen onomastische (namensbasierte) Erkennungsverfahren (Humpert/Schneiderheinze 2000, 2002) Abhilfe, weil viele Aussiedler russische Vor- und Familiennamen tragen. So deuten Namen wie Juri Idimetschenko recht eindeutig auf russische Herkunft hin. Doch schon bei Viktoria Kamrowskow treten Zweifel auf, und einen Helmut Kühnemeier oder Andreas Altstether kann kein Erkennungsalgorithmus als Aussiedler ausweisen.⁶ Die Methode leidet unter dem in § 94 des Bundesvertriebenengesetzes festgelegten Recht der Aussiedler, deutschsprachige Formen ihres Vor- und Familiennamens und ggf. sogar neue Vornamen anzunehmen. Eine hohe Ausschöpfung wird bei dieser Methode möglicherweise mit hohem Beifang (falsch positive Erkennung) nichtdeutscher osteuropäischer Personennamen erkaufte, die nicht zu Aussiedlern gehören, während eine hohe Treffsicherheit nur bei geringer Ausschöpfung erreichbar ist. Trotzdem könnte das namensbasierte Verfahren in Verbindung mit Screenerfragen dann für eine einfache und günstige Stichprobenziehung etwa aus dem Telefonbuch gut geeignet sein, wenn die Ausschöpfung kein Kriterium und gleichzeitig gewährleistet ist, dass die Namenscharakteristik als Determinante der Ziehungswahrscheinlichkeit nicht mit sozial-kulturellen Eigenschaften oder anderen Kriteriumsmerkmalen korreliert. Diesen Fragen wendet sich Abschnitt 3 zu.
5. Für Jugendlichen- und Schülerbefragungen wurden häufig Schulen bzw. Schulklassen als Auswahlrahmen benutzt (z. B. Strobl/Kühnel 2000; Babka v. Gostomski 2001; Wetzels et al. 2001). Dabei dienen Schulstatistiken über Aussiedleranteile zur Steuerung der Auswahl von Schulen. Einrichtungen mit geringem Anteil werden aus forschungsökonomischen Erwägungen häufig unterrepräsentiert, und die Präzision der auf Lehrereinschätzungen beruhenden Hintergrunddaten ist, wie allgemein bei prozessproduzierten Daten, manchmal kaum klar zu beurteilen. Da dieses Verfahren sich nicht auf die Aussiedlerpopulation übertragen lässt, sollen weitere Einzelheiten hier übersprungen werden.

In diesem Beitrag wird zunächst ein alternatives Verfahren vorgestellt, das durch Nutzung der Staatsangehörigkeiten und des Geburtsorts im Melderegister eine hohe Ausschöpfung bei geringem Beifang realisiert. Wegen der Nutzung von Ortsnamen (Toponymen) ist hier auch von einer toponomastischen Methode die Rede. Das Melderegister erwies sich bei einer ausführlichen Erörterung denkbarer Zuwanderer-Auswahlgesamtheiten (Salentin 1999) als beste verfügbare

6 Es handelt sich um fiktive Namen, die weder in den verwendeten Stichproben noch im aktuellen Telefonbuch der Bundesrepublik existieren. Jede Ähnlichkeit mit realen Personen wäre rein zufällig.

Grundlage. Das Verfahren ist aber auch auf andere Datenbestände anwendbar, sofern Geburtsorte zugänglich sind. Das Bundesamt für Migration und Flüchtlinge hat für eine Aussiedlerbefragung 2006 teilweise Geburtsorte als Kriterium zur Auswahl aus Melderegistern verwendet (Haug/Sauer 2007).⁷

2 Toponomastische Stichprobenziehung

2.1 Grundlagen

Weil sie als Deutsche im Sinn des Grundgesetzes weitgehend beim Grenzübertritt ihren deutschen Pass erhalten, gehen Aussiedler einerseits zu keinem Zeitpunkt in die amtliche Ausländerstatistik ein. Eine Ausnahme bilden nichtdeutsche Ehegatten der Deutschstämmigen, die mit diesen seit weniger als drei Jahren verheiratet sind, sowie sonstige mitreisende Familienangehörige. Sie bleiben Ausländer und müssen sich einbürgern lassen; über ihre Zahl existieren keine verlässlichen Angaben. Als deutsche Staatsangehörige sind Aussiedler wie in den meisten Statistiken andererseits scheinbar auch in Melderegistern nicht zu identifizieren. Das Melderegister bietet jedoch bisher unbeachtete Identifikationsmöglichkeiten unter Nutzung des Merkmals der zweiten Staatsangehörigkeit und des Geburtsorts. Obwohl bereits vor der Reform des Staatsangehörigkeitsrechts im Jahr 1999 die Bundesrepublik dem Prinzip der Vermeidung von Mehrstaatlichkeit folgte, wurden erhebliche Anteile der Gruppe der Aussiedler unter Hinnahme doppelter Staatsangehörigkeit eingebürgert. Nach Auskunft des Bundesministeriums des Inneren liegen dazu keine verlässlichen Zahlen vor, doch weil die Beibehaltung oder Wiederannahme der früheren Staatsangehörigkeit für Aussiedler bis ins Jahr 1999 rechtlich nicht zu beanstanden war, ist offensichtlich ein beträchtlicher Teil der Aussiedler damals Mehrstaater geworden. (Dies ist auch einer Pressemitteilung des früheren Innenministers Schily vom 4. August 2002 zu entnehmen, der zufolge zwischen 1993 und 2000 1,2 Millionen Aussiedler die deutsche Staatsangehörigkeit erhalten haben, ohne ihre frühere aufgeben zu müssen).⁸ 1994 besaßen im Aussiedlerpanel des Bundesinstituts für Bevölkerungsforschung drei Jahre nach Einreise 63% der Haushaltsvorstände aus Polen, Rumänien und der ehemaligen Sowjetunion neben der deutschen eine weitere Staatsangehörigkeit (Mamme/Schiener 1998: 37). Zum Anteil der Mehrstaater unter den Aussiedlern liegen amtlicherseits auch für die Zeit nach 2000 keine Publikationen vor. Selbst, wenn er wegen der geänderten Gesetzeslage zurückgegangen ist, dürfte gegenwärtig eine beträchtliche Zahl der Aussiedler an der Kombination bestimmter Staatsangehörigkeiten zu erkennen sein, etwa Bundesrepublik und Russische Föderation, Bundesrepublik und Kasachstan etc.⁹ In Teil 3.3 erfolgt eine Schätzung des Anteils.

7 Der Autor dankt Frau Sonja Haug für die Überlassung eines Vortragsmanuskripts und mündliche Informationen. Leider ist ein ausführlicher Forschungsbericht zur Studie bis zum Ende des MDA-Begutachtungsverfahrens nicht publiziert.

8 Bundesministerium des Inneren, http://www.bmi.bund.de/cln_028/nn_662928/Internet/Content/Nachrichten/Archiv/Pressemitteilungen/2002/08/Schily_zum_modernen_Id_89621_de.html (16.2.2006).

9 Während seit der Novelle des Staatsangehörigkeitsrechts die *Annahme* einer fremden Staatsangehörigkeit für Deutsche rechtswidrig ist, ist die *Beibehaltung* der früheren Staatsangehörigkeit, die den meisten Ausländern bei der Einbürgerung verweigert wird, bei Aussiedlern nicht zu beanstanden.

Es verbleibt aber ein Personenkreis, der seine frühere Staatsangehörigkeit aufgegeben hat, also nur die deutsche Staatsangehörigkeit besitzt, die sog. Einstaater. Er ist wiederum an Geburtsorten in Herkunftsgebieten der Aussiedler zu erkennen und kann anhand dieser Geburtsorte identifiziert werden, weil nach dem Melderechtsrahmengesetz des Bundes und den Meldegesetzen der Länder Geburtsorte zu den zulässigen Merkmalen einer sog. Melderegister-Gruppenauskunft, also einer Stichprobe, gehören. Die Auswahl nach Geburtsort ist allerdings technisch aufwendiger als eine statische Filterung nach einem Merkmal wie dem Alter oder der Staatsangehörigkeit, weil keine simple Regel wie Staatsangehörigkeit = Russland zugrunde gelegt werden kann. Es sind Schreibfehler und -varianten (Nowoneschinka, Nowoneshenka, Nowoneshinka), heute ungebrauchliche deutsche Ortsnamen (Grünfeld in Kirgistan, Alexanderhöh in Russland) sowie Verwaltungsgliederungszusätze (Pobotschnoje Geb. Saratow/Kasachstan) zu beachten. Ohne hier im Detail die auftretenden technischen Probleme abhandeln zu können, muss festgestellt werden, dass ein gewisser Programmieraufwand entsteht. Der Aufwand für die Zusammenstellung von Ortsnamenlisten ist so lange überschaubar, wie keine vollständige Ausschöpfung erzielt werden muss, weil bereits wenige Geburtsorte eine große Personenzahl identifizieren (s. u.). Mit entsprechendem Rechercheeinsatz kann aber auch ein solches Maximalziel weitgehend realisiert werden. Dabei ist es nicht notwendig, Synonyme und Schreibvarianten zusammenzufassen, jedoch muss eine möglichst eindeutige Zuordnung eines Orts zu einem Gebiet sichergestellt werden. Dazu gehört der Zugriff auf internationale Ortsdatenbanken.¹⁰

2.2 Ausschöpfung der Grundgesamtheit

Identifikation der Geburtsorte

Anhand eines Melderegisterauszugs wird im Folgenden abgeschätzt, zu welchem Grad das Verfahren in der Lage ist, Aussiedler zu identifizieren, ohne dabei fälschlicherweise Nicht-Aussiedler zu entdecken. Eine *exakte* Aussage zur Ausschöpfung wird kaum jemals möglich sein, weil der Umfang der Einstaater innerhalb der lokalen Aussiedler-Grundgesamtheit unbekannt ist, es sein denn, man könnte die Geburtsorte aller Deutschen einschließlich der aus der ganzen Welt stammenden Eingebürgerten lückenlos bestimmen. Die Größenordnung der erreichbare Ausschöpfung lässt sich aber umreißen. Hier geschieht dies in einer Pilotanwendung der Toponomastik auf Mehrstaater in Verbindung mit der publizierten Migrationsstatistik. Zunächst liegt uns ein Bestand 90.000 bundesdeutscher, 733 russischer, 697 kasachischer, 126 kirgisischer, 1136 rumänischer und 2.435 polnischer Ortsnamen (einschließlich der Schreibvarianten und älterer Bezeichnungen) vor. Untersucht wurden damit die anonymen Daten der über 20.000 Bielefelder Einwohner, die die erste oder zweite Staatsangehörigkeit Russlands, Kasachstans, Kirgisistan, Polens oder Rumäniens,

10 Z. B. der GEOnet Names Server der National Geospatial-Intelligence Agency (NGA) der USA, nationale Statistikbehörden und Postverwaltungen.

also der Hauptherkunftsländer der Aussiedler¹¹, besitzen, mit dem Ziel, nur durch den Geburtsort die Herkunftsländer zu identifizieren. Die Geburtsorte gehören ausnahmslos zu Ausländern (15,8%) oder Mehrstaatern (84,2%¹²). Die folgende Tabelle 1 weist die durch Datenbankanalyse maschinell ermittelten Geburtsländer nach aktueller Nationalität aus. Die Nationalitäten fassen die Erststaatsangehörigkeit der Ausländer und die Zweitstaatsangehörigkeit der deutschen Doppelstaater zusammen. (Letztere besitzen als erste Staatsangehörigkeit immer die deutsche.) An dieser Stelle ist es nicht möglich, auch deutsche Einstaater mit Geburtsort in einem Herkunftsland wie Kasachstan – also mutmaßliche Aussiedler – zu ermitteln. Dies folgt in Abschnitt 2.3 auf Grundlage eines zweiten Datensatzes.

Tabelle 1 Staatsangehörigkeiten und identifizierte Geburtsorte

	Ausländische Staatsangehörigkeit										Gesamt
	Polen		Rumänien		Russland		Kasachstan		Kirgisistan		
	N	%	N	%	N	%	N	%	N	%	
Geburtsort											
BRD	1808	20,5	71	13,2	836	13,7	536	13,4	246	19,5	3497
Polen	5092	57,7	0	0,0	3	0,0	2	0,0	1	0,1	5098
Rumänien	1	0,0	410	76,4	0	0,0	0	0,0	0	0,0	411
Russland	2	0,0	0	0,0	2568	42,1	264	6,6	131	10,4	2965
Kasachstan	0	0,0	0	0,0	515	8,4	1985	49,5	75	5,9	2575
Kirgisistan	3	0,0	0	0,0	130	2,1	41	1,0	516	40,8	690
Sonstige	40	0,5	1	0,2	154	2,5	79	2,0	37	2,9	311
uneindeutig	1125	12,7	53	9,9	397	6,5	212	5,3	81	6,4	1868
unbekannt	759	8,6	2	0,4	1501	24,6	894	22,3	177	14,0	3333
Gesamt	8830	100	537	100	6104	100	4013	100	1264	100	20748

Quelle: Melderegister der Stadt Bielefeld, Stand 13.12.2005; eigene Analysen.

Es zeigen sich Unterschiede, aber auch Gemeinsamkeiten zwischen den Herkunftsländern. Für 57,7% der polnischen Staatsangehörigen (einschließlich der Deutsch-Polen) ließ sich ein Geburtsort in Polen ermitteln. 20,5% sind in der Bundesrepublik geboren (darauf wird gleich zurückzukommen sein). 12,7% der Orte lassen sich keinem Land zuordnen, weil ihr Name in mehreren Ländern auftritt. Ein Beispiel für diesen Fall ist „Eichenau“, das es sowohl im Landkreis Fürstfeldbruck in Bayern als auch als alte Bezeichnung für Katowice-Mala Dabrowka (Woiwodschaft Slaskie/Oberschlesien) in Polen gibt. 8,6% der Geburtsorte sind bisher nicht ländercodiert. Die Geburtsorte der anderen Staatsangehörigkeiten sind seltener uneindeutig, weil in den betreffenden Ländern deut-

11 Aus den UdSSR-Nachfolgestaaten, Rumänien und Polen stammen 94% aller Aussiedler (Quelle: s. Fußnote 1). Die Ukraine als wichtiges aus der Sowjetunion hervorgegangenes Herkunftsländ bleibt hier unberücksichtigt.
12 Anteil der Mehrstaater: Polen 80,8%; Rumänien 67,8%; Russland 85,1%; Kasachstan 91,0%, Kirgisistan 88,9%.

sche Bezeichnungen die Ausnahme bilden. Ortsnamen treten aber oft mehrfach in Russland, der Ukraine, Kasachstan und anderen Staaten auf. Fasst man die Nachfolgestaaten der Sowjetunion zusammen, sinkt der Anteil uneindeutiger Zuordnungen unter 1%. Gleichzeitig sind die Geburtsorte der russischen, kasachischen und kirgisischen Staatsangehörigen häufiger unidentifiziert, was zum einen an den vielfältigen Varianten der ursprünglich kyrillisch geschriebenen Namen in lateinischer Schrift und zum anderen an der im Vergleich mit den ehemaligen deutschen Siedlungsgebieten in Polen, besonders in Schlesien, wesentlich schlechteren Dokumentation der Migrationsgeschichte liegt.¹³ Auffällig zeigt sich die Verschlingung der (bekanntlich nicht immer freiwilligen) Wanderungsverläufe der Deutschstämmigen innerhalb der Sowjetunion: 10,4% der kirgisischen Staatsangehörigen sind in Russland geboren, während von den 2.575 in Kasachstan Geborenen 264 Personen (6,6%) aktuell einen Pass der russischen Föderation besitzen.

Die Nachkommen der Aussiedler der ersten Generation sind selbstverständlich nicht an ihrem Geburtsort zu erkennen, der ja definitionsgemäß im Inland liegt. Dagegen haben offenbar auch viele Kinder der Aussiedler mehrfache Staatsangehörigkeiten inne, wie die Analyse der Geburtsorte ergibt. Von den in Bielefeld gemeldeten Staatsangehörigen der Russischen Föderation (Ein- und Doppelstaater) sind 14% in der Bundesrepublik geboren. Dieser Effekt kommt nicht durch fehlerhafte Ortsnamenerkennung zustande, denn alleine 12% geben Bielefeld als Geburtsort an. Die Anteile sind für die anderen Staatsangehörigkeiten teilweise sogar noch deutlich größer. Es ist damit sichergestellt, dass zumindest auf absehbare Zeit Aussiedler-Stichprobenziehungen nach Staatsangehörigkeit auch für die Folgegeneration möglich sein werden, wenngleich keine Aufschlüsse darüber vorliegen, welche Ausschöpfung der zweiten Generation auf diese Weise erreicht wird und welche strukturellen Unterschiede zwischen Ein- und Mehrstaaten bestehen.

Die theoretische Obergrenze der erzielbaren Erkennungsleistung ist erreicht, wenn ein vollständiges Ortsverzeichnis zugrunde gelegt werden kann. Realistisch ist es dagegen anzunehmen, dass die Lage der Hälfte der zur Zeit unbekannten Orte mit vertretbarem Aufwand noch recherchiert werden kann.¹⁴ Damit lässt sich aus Tabelle 1 errechnen: 95,7% der Geburtsorte in Polen, 87,7% der Orte in Russland etc. und über die Staaten gewichtet gemittelt 92,0% aller Aussiedler-Geburtsorte lassen sich lokalisieren. Wenn nun etwa die Hälfte der Aussiedler bereits als Doppelstaater zu erkennen ist (dazu Abschnitt 2.3) und von der anderen Hälfte, den deutschen Einstaaten, 92% mit dem Geburtsortsverfahren ermittelt werden können, lässt sich durch Kombination von Staatsangehörigkeit und Geburtsort eine Erkennungsrate von gut 95% aller Aussiedler realisieren.

13 Zum früheren deutschen Verwaltungswesen in Schlesien und persönlichen Biographien finden sich zahlreiche Druckschriften und Internetseiten.

14 Hilfreich wäre die Einbeziehung der 101.418 Orte umfassenden Liste typischer Aussiedler-Geburtsorte des Bayrischen Landeskriminalamtes, die das Bundesamt für Migration und Flüchtlinge nutzte (Haug/Sauer 2007: 8).

Falsch positive Identifikation (Beifang)

Welche Personen erkennt das Verfahren fälschlicherweise? Diese Frage interessiert vornehmlich aus ökonomischen Gründen, nämlich im Hinblick auf überflüssige Kontakte, denn letztlich lässt sich ja eine Beeinträchtigung der realisierten Stichprobe durch Zielgruppenfremde mit Screenerfragen leicht vermeiden. Doppelstaater, die gleichzeitig Deutsche und Angehörige eines Herkunftslandes von Aussiedlern sind, ohne selbst deutscher Abstammung zu sein, dürften aufgrund der restriktiven deutschen Einbürgerungsrechts quantitativ unbedeutend bleiben. Eher stellen sonstige Personen ein Problem dar, die in der UdSSR oder ihren Nachfolgestaaten, Polen und Rumänien geboren sind. Hier ist zwischen Ausländern und deutschen Einstaatern zu unterscheiden.

Ausländer stellen nur insofern ein Problem dar, als der vollständige Verzicht auf sie den Aus-schluß der oben erwähnten Aussiedler-Haushaltsangehörigen (frisch verheiratete Ehegatten, sonstige Mitreisende) nach sich zieht. Am Beispiel der russischen Nationalität kann man ablesen, dass in diesem – unter inhaltlichen Gesichtspunkten jedoch nicht sonderlich relevanten – Fall erheblicher Beifang zu erwarten ist. Am 31.12.2004 lebten knapp 179.000 Personen russischer Nationalität in der Bundesrepublik.¹⁵ Darunter sind neben ehemaligen Angehörigen der Roten Armee etc. hauptsächlich jüdische Kontingentflüchtlinge aus den GUS-Staaten. Die Zahl letzterer liegt zwischen den 219.604 für die Gruppe registrierten Einreisen und Aufnahmezusagen (bis Ende 2004) und den 111.811 am 18.4.2005 im Ausländerzentralregister (AZR) als Kontingentflüchtlinge verzeichneten GUS-Staatsangehörigen (Haug/Schimany 2005: 8). Die genaue Zahl ist unbekannt, weil viele Juden in Drittstaaten weiterreisen; viele inzwischen Eingebürgerte bleiben dem AZR verborgen. Offensichtlich ist von deutlich mehr Personen ohne als mit Beziehung zu einem Aussiedler-Haushalt auszugehen.

Etwas schwieriger ist der Beifang unter Eingebürgerten zu quantifizieren. Die Eingebürgertenstatistik enthält seit 2000 keine Aussiedler mehr. Für frühere Zeiträume ist nach Auskunft des Bundesministeriums des Inneren¹⁶ zwar nicht sichergestellt, dass alle Aussiedler enthalten sind, aber zumindest ein großer Teil, was die Bestimmung des Umfangs nichtdeutscher Einbürgerungen erschwert. Eine Schätzung ist dennoch möglich. Bis 1988 wurden den Statistischen Jahrbüchern zufolge jährlich maximal 1.000 Personen aus der UdSSR eingebürgert. Seit 2000 sind es im Mittel 4.100 jährlich. Dagegen waren es von 1989 bis 1999 im Durchschnitt 118.800 pro Jahr. Offensichtlich erklärt sich die große Zahl in dieser Periode nur durch die enthaltenen Aussiedler. Nehmen wir nun an, dass seit 1985 im Mittel nur ca. 4.000 Personen nichtdeutscher Abstammung jährlich die deutsche Staatsbürgerschaft verliehen wurde – davor fanden kaum Einbürgerungen statt – und alle diese Personen noch im Bundesgebiet leben. (Der Versuch einer weitergehenden Quantifizierung der Abwanderung für diesen Personenkreis wie auch für Aussiedler führt zu weit, siehe aber die Überlegungen bei Mamme/Schiener (1996) und Engels (2000: 2).) Dann stünden 2,3 Mio. Aussiedlern aus der ehemaligen UdSSR ca. 75.000 sonstige Eingebürgerte gegenüber, und eine Beifangquote von 3% wäre folglich zu erwarten.

15 Statistisches Bundesamt, <http://www.destatis.de/basis/d/bevoe/bevoetab10.php> (3.7.2007).

16 Telefonische Auskunft der Abteilung M 2.5.

Ein ähnlicher Überschlag ergibt für polnischstämmige Aussiedler bei ca. 80.000 Eingebürgerten einen Beifang von 5,2%. Wegen geschichtlicher Besonderheiten stellt Polen allerdings in zweifacher Hinsicht einen Sonderfall dar, weshalb diese Quote stark irreführend ist. Eine erste, nicht quantifizierbare, aber vermutlich kleine Störgröße bilden in diesem Fall Personen, die in den ehemaligen deutschen Ostgebieten im heutigen Polen geboren wurden, ihren Wohnsitz aber in der Vorkriegszeit als Binnenmigranten im Deutschen Reich spontan gen Westen verlassen und nicht erst, wie es §1 (1) des Bundesvertriebenengesetzes zur Erfüllung der Vertriebeneneigenschaft fordert, „im Zusammenhang mit den Ereignissen des zweiten Weltkrieges infolge Vertreibung, insbesondere durch Ausweisung oder Flucht, verloren“ haben. Zweitens sind in der unmittelbaren Nachkriegszeit mehrere Millionen Vertriebene ohne Einbürgerung in die Bundesrepublik und in die Deutsche Demokratische Republik gelangt (Faulenbach 2002: 44). Dort fanden sie u. a. wegen der vergleichbaren Vorkriegssituation und des Lastenausgleichsgesetzes Startvoraussetzungen vor, die denen der alteingesessenen Bevölkerung nicht unähnlich waren. Es wäre unangemessen, sie zugleich mit den Aussiedlern ab den achtziger Jahren in eine Stichprobe zu ziehen, und ein aufwendiges Screening wird daher unvermeidlich. Wegen der eher geringen Zahl der zwischen 1960 und 1986 eingereisten Personen¹⁷ nehmen sich in den Geburtskohorten ab 1960 allerdings die Effizienzprobleme weniger dramatisch aus als in früheren. Nichtsdestoweniger legt diese Überlegung es nahe, dem uneingeschränkt funktionierenden Doppelstaaterprinzip für Aussiedler aus Polen der Vorzug gegenüber dem Geburtsortsprinzip zu verleihen.

2.3 Ist die Toponomastik verzichtbar?

Das führt zu der Frage, ob das Kriterium der Doppelstaatigkeit für eine Stichprobenziehung aus dem Melderegister ausreicht bzw. inwiefern auf die ja viel aufwendigere Auswertung des Geburtsorts durch das toponomastische Erkennungsverfahren verzichtet werden kann. Dies wäre allenfalls bei einem kleinen Anteil deutscher Einstaater der Fall. Um diese Frage zu beantworten, wurde ein zweiter Datensatz herangezogen. Er enthält die Geburtsorte und Staatsangehörigkeiten aller nicht in Bielefeld selbst geborenen Personen im Bielefelder Melderegister, also auch die Geburtsorte aller Deutschen, die in Mannheim, New York, Timbuktu oder Omsk geboren sind.¹⁸ Diese Datenbasis schließt alle Aussiedler im juristischen Sinn ein, auch aktuell deutsche Einstaater, da diese im Ausland geboren sein müssen. Die eindeutig in den wichtigsten Herkunftsstaaten der Aussiedler geborenen Personen wurden identifiziert. Uneindeutige Gebietszuordnungen und gänzlich unbekannte Geburtsorte bleiben nun unberücksichtigt. Für die Proportionen der Staatsangehörigkeiten dürfte dies aber nicht von erheblicher Bedeutung sein.

Die Zahl der möglichen Kombinationen von Geburtsländern und Staatsangehörigkeiten ist unübersichtlich. So ist es möglich, dass in einem Ort im heutigen Kasachstan geborene Deutsch-

17 Vgl. Fußnote 2.

18 187.840 von 331.151 Bielefeldern sind außerhalb der heutigen Stadtgrenzen geboren (56,7%).

stämmige einen russischen Pass besitzen, z. B. weil sie noch zu Sowjetzeiten den Wohnsitz gewechselt haben. Zur Vereinfachung wurde eine Zusammenfassung zu der technischen Variablen *Staatigkeit* vorgenommen (siehe Tabelle 2). Staatigkeit bedeutet die Kombination von Staatsangehörigkeiten und hat hier drei Ausprägungen. Unter *Einstaatern* sind Personen mit nur einer Staatsangehörigkeit, hier der deutschen, zu verstehen. *Mehrstaater* besitzen mindestens, aber fast ausnahmslos genau, zwei Staatsangehörigkeiten, von denen hier immer eine die deutsche ist. Die zweite ist mehrheitlich, aber nicht ausschließlich, die des Geburtslandes. Alle Kombinationen wie die obige werden damit erfasst. Die Kategorie *Ausländer* definiert sich durch das Fehlen der deutschen Staatsangehörigkeit und umfasst sowohl ausländische Ein- als auch Mehrstaater (letztere etwa mit polnischem und gleichzeitig französischem Pass).

Tabelle 2 Staatigkeit nach Geburtsland

(1) Geburtsland	(2) dt. Ein- staater*	(3) % der Stichprobe	(4) % der Deutschen	(5) dt. Mehr- staater*	(6) % der Stichprobe	(7) % der Deutschen	(8) Ausländer*	(9) % der Stichprobe	(10) Summe
Russland	2.278	41,1	45,1	2.770	50,0	54,9	490	8,8	5.538
Polen	11.603	69,4	74,4	3.983	23,8	25,6	1.137	6,8	16.723
Kasachstan	2.785	49,6	52,5	2.516	44,8	47,5	310	5,5	5.611
Kirgisistan	1.054	57,1	61,0	674	36,5	39,0	119	6,4	1.847
Rumänien	274	39,9	51,8	255	37,2	48,2	157	22,9	686
Summe	17.994	59,2	63,8	10.198	33,5	36,2	2.213	7,3	30.405

Quelle: Melderegister der Stadt Bielefeld, Stand 17.3.2006; eigene Auswertungen.
* siehe Anmerkungen im Text.

Die Fallzahlen in Tabelle 2 sind zum einen auf die gesamte Stichprobe prozentuiert. Zunächst ist zu erkennen, dass der Anteil der Ausländer, abgesehen von Rumänien, nur im einstelligen Prozentbereich liegt (Spalte 9). Zum anderen ist für deutsche Ein- und Mehrstaater der Prozentanteil bezogen auf die Summe dieser beiden Ausprägungen angegeben (Spalten 4 u. 7). Sie entspricht abgesehen von sehr wenigen nicht eingebürgerten Aussiedlern sowie eingebürgerten Nicht-Aussiedlern der Zahl der Aussiedler (Ausnahme: Polen, s.o.). Russlanddeutsche Aussiedler sind demnach mehrheitlich Doppelstaater (54,9) während gut die Hälfte der kasachischen, kirgisischen und rumänischen Aussiedler nicht mehr die Staatsangehörigkeit des Herkunftslandes innehat (51,8% bis 61,0%, Spalte 4). Ob der Verzicht auf deutsche Einstaater zu qualitativen Verzerrungen führt, weil sie systematisch andere sozialstrukturelle Merkmale aufweisen als Doppelstaater, kann sicher letztlich nur substantielle Forschung zeigen. Zumindest bis zum Beweis des Gegenteils und auf jeden Fall unter dem Gesichtspunkt der quantitativen Ausschöpfung muss das Fazit aber lauten, dass das toponomastische Verfahren unerlässlich ist.

2.4 Administrative und juristische Implementationsvoraussetzungen, Kosten

Obwohl das toponomastische (wie auch das onomastische) Verfahren nicht an einen speziellen Auswahlrahmen gebunden ist, wird doch die Auswahl aus dem kommunalen Melderegister der häufigste Anwendungsfall sein. Die Mitwirkung der Meldeämter ist daher eine zentrale Implementationsvoraussetzung. Die juristischen Grundlagen sind insofern gegeben, als 1. zumindest Universitäten und vergleichbare Forschungseinrichtungen zu den „sonstigen öffentlichen Stellen“ zu zählen sind, die das Melderecht als Empfänger von Gruppenauskünften vorsieht, 2. die Daten zur Erfüllung der Forschungsaufgabe erforderlich sind, weil Aussiedler vorrangig am Geburtsort zu erkennen sind, und 3. die benötigten Merkmale in den Meldegesetzen auch explizit als mitteilbar genannt werden. Dass hier Listen von Geburtsorten als Input verwendet werden, dürfte, juristisch betrachtet, nicht von Bedeutung sein, da ja auch konventionelle Melderegisterstichproben auf – allerdings wesentlich kürzere – Listen von Staatsangehörigkeiten zurückgreifen; es ist zu bedenken, dass ja sogar Auskünfte über namentlich genannte Einzelpersonen melderechtlich unter bestimmten Voraussetzungen erlaubt sind. Allerdings wird man dies im Einzelfall Amtsleitern und örtlichen Datenschutzbeauftragten erläutern müssen. Ablehnungen des Verfahrens wurden auf sechs bisher gestartete Anfragen an Städte hin nicht erteilt. Dass zumindest simple melderegisterbasierte Stichproben bundesweit möglich sind, zeigt der ALLBUS (Haarmann et al. 2006).

Die technische Realisierung dürfte indes oft ein Problem darstellen, weil die Programmierung des Selektionsalgorithmus an die Grenzen der Möglichkeiten der datenführenden Dienststellen stößt. Auch die vereinfachte Version, der starre 1:1-Abgleich eines Bestands von Ortsnamen mit den lokal gespeicherten Geburtsorten, verursacht in Ämtern und beauftragten Rechenzentren nicht unerheblichen Programmieraufwand, da es sich um ein zwar keineswegs generell in der Datenbankprogrammierung, jedenfalls aber im Verwaltungsvollzug ungewöhnliches Verfahren handelt. Die Kosten für den Zusatzaufwand wird in der Regel der Forscher tragen müssen, weshalb das Verfahren in jedem Fall merklich teurer als herkömmliche Melderegisterstichproben sein muss. Leider ist an dieser Stelle keine allgemein gültige Kostenschätzung möglich, weil die Kalkulationsmodelle der Länder und Kommunen zu stark variieren. So sieht Hamburg einen Korridor von 100,- bis 12.500,- Euro je Gruppenauskunft vor, während die Gebührenordnung Niedersachsens eine Pauschale von 14,80 bis 35,70 Euro zuzüglich 3 bis 15 Cent je ausgewählten Einwohner und 0,1 bis 3 Cent je registrierten Einwohner festlegt.¹⁹ Für eine Stichprobe von N=2.000 bei einem Einwohnerbestand von 200.000 Personen bleiben die Kosten ohne Berücksichtigung des Programmieraufwands im günstigsten Fall unter 300, im ungünstigsten steigen sie über 6.300 Euro. Die Rechnungsmodelle der anderen Bundesländer sind ähnlich flexibel, wobei z. B. in Baden-Württemberg zumindest die Komponente für die gemeldeten, aber nicht ausgewählten Einwohner ent-

19 http://hh.juris.de/hh/MeldeGebO_HA_1993_P1.htm, http://cdl.niedersachsen.de/blob/images/C1736588_L20.pdf (18.4.2006). Die Gebührenordnungen zahlreicher Gebietskörperschaften sind im Internet verfügbar. Weitere Details würden den verfügbaren Rahmen sprengen.

fällt.²⁰ Die Kosten schwanken auch innerhalb eines Bundeslandes. Verwaltungen mancher Städte erlassen Universitäten die Gebühren oder rabattieren sie erheblich (wie Stuttgart). Realistisch erscheint dem Autor der Durchschnittsbetrag von 200 bis 1.000 Euro je Kommune, den Orte wie Dortmund, Stuttgart und Hamburg auf Anfrage nannten; leider sind auch Ausreißer von über 10.000 Euro bekannt.

Dies ist, neben der bisher herrschenden Unklarheit über Referenzparameter, sicher einer der Gründe, warum ein naheliegendes – wenn man will, eigentlich triviales – Verfahren in der Vergangenheit kaum eingesetzt worden ist. Der Ressourcenaufwand für Flächenstichproben ist hoch, allenfalls sind zweistufige Verfahren mit wenigen Kommunen auf einer ersten Auswahlstufe vorstellbar, wie sie das Bundesamt für Migration und Flüchtlinge (Haug/Sauer 2007: 5) implementierte. Für bundesweit repräsentative Erhebungen werden die Kosten in der jetzigen kommunalen Struktur des Meldewesens wohl immer prohibitiv bleiben.

3 Onomastische Stichprobenziehung

3.1 Numerische Beschreibung

Die Stichprobenziehung auf der Grundlage spezifischer Namen verdient eine genaue Prüfung, weil sie in Verbindung etwa mit dem Telefonbuch als Auswahlgrundlage eindeutige Vorzüge aufweist: Sie ist kostengünstig, leicht zu implementieren und problemlos auf beinahe beliebige Flächeneinheiten ausdehnbar. Im folgenden wird geprüft, inwiefern die in der Einleitung angedeuteten möglichen Nachteile diese Vorteile zunichte machen.

Als Grundlage dient eine im Juni 2004 für eine Studie zu Diskriminierungserfahrungen aus dem Einwohnermelderegister der Stadt Bielefeld gezogene Zufallsstichprobe von 1732 Aussiedlern (neben anderen wichtigen Herkunftsgruppen) im Alter von 18 bis 25 Jahren. Ziehungskriterium war für sie ein Geburtsort in einem Herkunftsland von Aussiedlern. Die Verteilung nach Geburtsland (Russland 47,9%, Kasachstan 33,8%, Kirgisistan 11,3%, Polen 6,8%, sonstige Länder 0,3%) weist ein leichtes Übergewicht aus Russland stammender Personen auf, weil bei der Stichprobenziehung ein sehr frühes Ortsregister mit niedriger Ausschöpfung der anderen Herkunftsländer zum Einsatz kam. Es ist auch nicht auszuschließen, dass in der Bruttostichprobe, für die nur teilweise Befragungsergebnisse zur Validierung vorliegen, auch eine kleine Zahl von Nicht-Aussiedlern, v. a. aus Polen, enthalten ist.

Die Vor- und Nachnamen in der Stichprobe wurden nach ihrer Typik in die Kategorien 1. *deutsch*, 2. *für das Geburtsland typisch* sowie 3. *unbekannt oder für ein Drittland typisch* eingeordnet. Dabei konnte nur auf Allgemeinwissen zurückgegriffen werden, weil keine elabo-

20 Der Einfluß der Einwohnerzahl auf die Kosten einer Meldeamtsstichprobe hat im ALLBUS zu einer Anpassung des Stichprobendesigns geführt, vgl. Haarmann et al. 2006: 55, FN. 21.

rierte Quelle wie die von Huschka/Gerhards/Wagner (2005) genutzte Personennamen-Datei der Personennamen-Beratungsstelle der Universität Leipzig zur Verfügung stand. Da die Daten aus dem Melderegister stammen, soll im Interesse der Vertraulichkeit die Zuordnung nicht im Detail aufgeführt werden. Oft enthalten herkunftslandtypische fremde Varianten der Namen die Zeichenfolgen *ks* anstelle von *x*, *ei* oder *ej* anstelle von langem *e* oder *é*, *dr* statt *der*, *z* oder *sz* statt *s*, *ija* oder *ja* anstelle von *ia* etc. Vor allem wurden in der Bundesrepublik nur selten vorkommende Namen russischen oder polnischen Ursprungs als typisch für Russland, Polen etc. klassifiziert. Zwischen den Ländern unterscheiden sich die Verteilungen der Namen-Typik nicht sehr stark. Lediglich Personen aus Polen tragen häufiger Nachnamen des Herkunftslandes sowie andererseits deutsche Vornamen als der Rest der Stichprobe. Eine Kreuztabelle von Vor- und Nachnamen der Gesamtstichprobe (Tabelle 3) zeigt, dass deutsche Namen eindeutig überwiegen (63,7% der Vor-, 77,8% der Nachnamen). Mehr als die Hälfte (52,4% der Personen) trägt sowohl einen deutschen Vor- als auch Nachnamen.

Tabelle 3 Namenstypik der Aussiedler

	Nachname						Summe	
	Deutsch		Herkunftsland		Unklar		N	%
	N	%	N	%	N	%		
Vorname								
Deutsch	907	52,4	161	9,3	35	2,0	1103	63,7
Herkunftsland	419	24,2	151	8,7	22	1,3	592	34,2
Unklar	22	1,3	12	0,7	3	0,2	37	2,1
Summe	1348	77,8	324	18,7	60	3,5	1732	100,0

Quelle: Melderegister der Stadt Bielefeld, Juni 2004; eigene Analysen.

Die Verteilung der Nachnamen ist deutlich flacher als die der Vornamen. Es wurden einschließlich der Schreibvarianten 1112 verschiedene Nach-, aber nur 382 Vornamen ermittelt. Deutsche Vornamen kommen durchschnittlich 5,0 mal, fremde 4,4 mal vor, deutsche Nachnamen 1,8 mal, fremde 1,1 mal. (Wegen der unterschiedlichen Besetzung der Herkunfts-Kategorien sind diese Werte nicht direkt vergleichbar.) Die Hitliste der häufigsten Vornamen wird von Al. (96 Vorkommen), Vi. (64), An. (57) bzw. den weiblichen Namen Ol. (48), Ta. (39) und An. (37) angeführt. In der Menge der 38 in der Stichprobe mehr als zehnmal vorkommenden Vornamen befinden sich 26 für den deutschen Sprachraum allgemein typische. Somit ist abzusehen, dass eine Stichprobe auf Vornamenbasis nicht unerheblichen Beifang erzielen würde. Da sich lokale Clusterungen, etwa durch Kettenmigration, kaum auf die Vor-, aber nachhaltig auf die Nachnamenverteilung auswirken, soll letztere hier nicht aufgefächert werden. Es muss genügen zu sagen, dass sich unter den 16 häufigsten die im deutschen Sprachraum allgemein sehr verbreiteten Familiennamen Schm. (231.000 mal im Telefonbuch), Br. (52.000 mal), Pe. (38.000 mal), Schr. (63.000 mal), Wo. (62.000 mal) und Be. (91.000 mal) befinden.

3.2 Abschätzung der Trennschärfe

Für eine grobe Abschätzung der onomastisch erzielbaren Trennschärfe wurden die Namen der Stichprobe mit dem Bestand der privaten Anschlüsse in einem maschinenlesbaren Telefonbuch verglichen. Idealerweise müssten die Aussiedlernamen einem bundesweiten Bestand (nicht mit einer Bielefelder Stichprobe, weil dabei örtliche Besonderheiten zusätzliche Probleme bereiten würden) aus der Zeit vor Beginn der Einreisewelle aus den deutsch besiedelten Gebieten in Mitteleuropa und der ehemaligen Sowjetunion gegenübergestellt werden, was aber technisch nicht zu realisieren ist. Daher wurde eine frühe Telefon-CD aus dem Jahr 1997 benutzt. Sie enthält 33 Mio. private Anschlüsse. 3 Mio. Vornamen sind nicht vorhanden oder abgekürzt. Im Jahr 1997 war allerdings ein erheblicher Teil der Aussiedler schon eingereist, wobei hingegen ihre Neigung, sich auch ins Telefonbuch eintragen zu lassen, unbekannt ist. Man kann annehmen, dass durch die zeitweilige Heimunterbringung viele Aussiedler erst mit Verzögerung in die Telefonbücher eingehen. Es kommt bei dieser Vorgehensweise durch einige Unbekannte sicher zu Schätzfehlern, die aber hingenommen werden können, da nur eine Größenordnung in Erfahrung gebracht werden soll. Ein einfaches Maß der Trennschärfe ist das Verhältnis der Personenzahl eines gegebenen Namens in der Stichprobe zur Zahl der auf einen solchen Namen gemeldeten Telefonanschlüsse. Die zentrale Strategie der Onomastik besteht darin, bevorzugt Namen mit einem hohen Quotienten zu verwenden wie Ir., der sich 36 mal in der Stichprobe und fast 5.500 mal im Telefonbuch befindet, und nicht etwa Pe., den es zwar auch 27 mal in der Aussiedlerstichprobe, aber vor allem 532.000 mal im Telefonbuch gibt. Daher wurden alle Namen nach dem Trennschärfequotienten sortiert. Dann wurde ermittelt, wie sich mit schlechter werdender Trennschärfe die Zahl gefundener Telefonanschlüsse verändert (dazu Abbildungen 1 und 2).

In Abbildung 1 ist gegen die Ausschöpfung der Bielefelder Stichprobe (x-Achse) abgetragen, wie viele Telefonanschlüsse mit den für die jeweilige Ausschöpfung notwendigen Vornamen gefunden werden (gestrichelte Linie, linke y-Achse). Wie man sieht, ist das Verfahren für eine Ausschöpfung bis ca. 60% relativ effizient. Die Vornamen von 60% der Stichprobe existieren 441.000 mal im Telefonbuch. Bei einer angenommenen Haushaltsgröße von durchschnittlich 2,5 Personen und einer Telefonbuch-Eintragsquote von 33% (beides grobe Schätzungen) würden somit 3,3 Mio. Personen erreicht. Dies entspricht der Größenordnung der Zahl aller zwischen 1950 und Ende 1996 eingereisten Aussiedler (3,2 Mio.), unter denen Personen aus der ehemaligen UdSSR und aus Polen ja den weitaus größten Anteil haben. Zweifelsohne ist es unrealistisch, nun anzunehmen, mit dem Verfahren würde eine Trennschärfe von annähernd 100% erzielt, weil die Zahlen fast gleichauf liegen. Doch andererseits kann man schließen, dass selbst bei Annahme ungünstigerer Parameter (Stichprobeneffekte, geringere Eintragsdichte der Aussiedler etc.) die Trennschärfe nicht maßlos sinkt, was den Screening-Aufwand prohibitiv in die Höhe schnellen ließe. Abbildung 1 zeigt allerdings auch, dass hauptsächlich fremde Namen durch hohe Trennschärfe hervorgehoben sind (obere Linie, rechte y-Achse). Einige seltene deutsche Doppelnamen erweisen sich als hochgradig trennscharf, weshalb der Anteil deutscher Namen anfangs noch eher hoch ist und dann fällt.

Abbildung 1 Trennschärfe der Vornamen

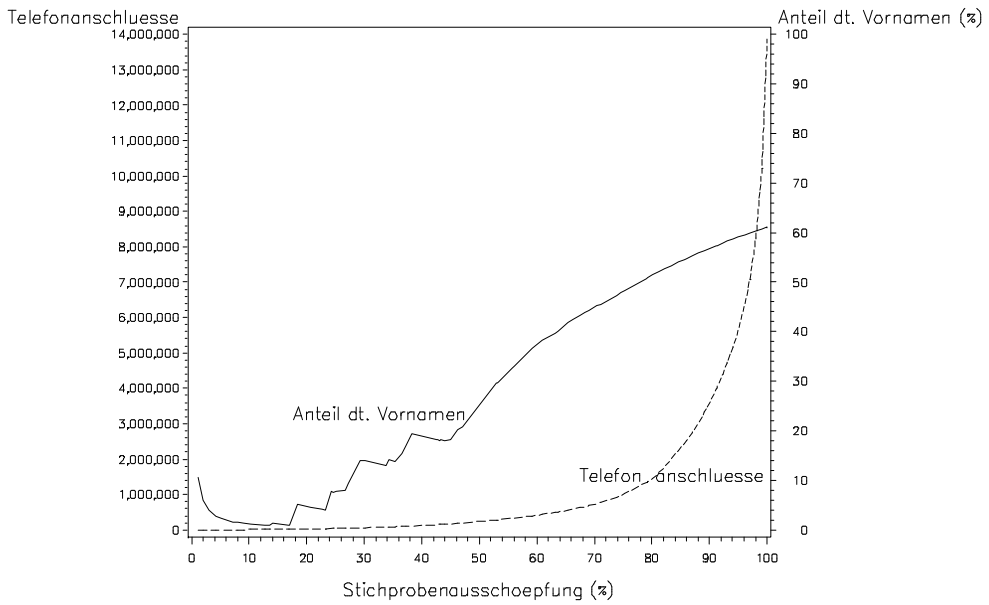
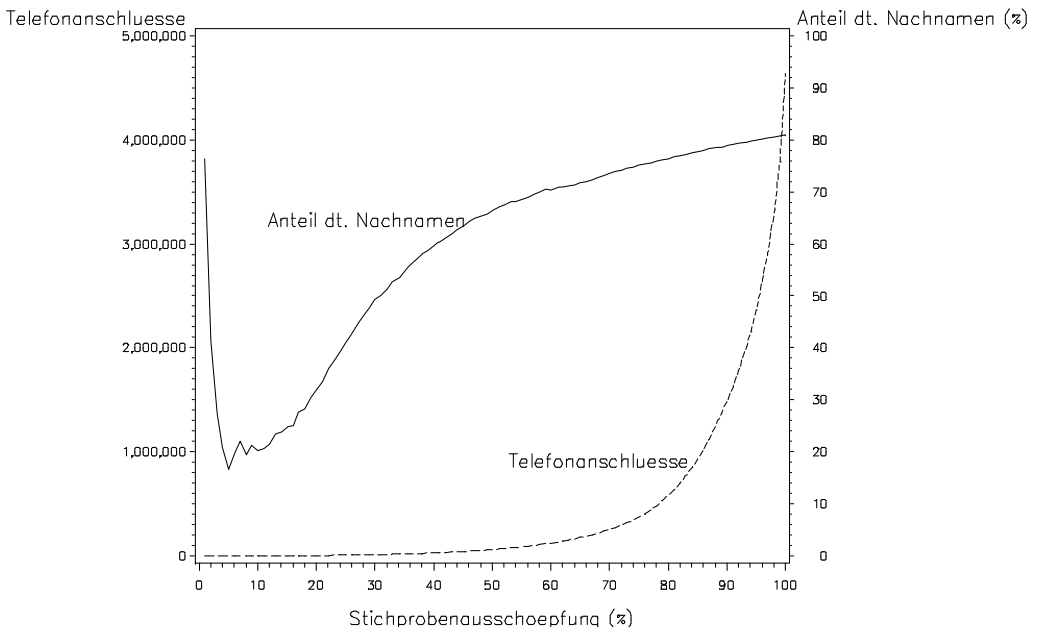


Abbildung 2 Trennschärfe der Nachnamen



Bei 60%iger Ausschöpfung liegt der Anteil deutscher Namen bei nur 38,2%. Um keine Disparität entstehen zu lassen, müsste also eigentlich die Ausschöpfung erhöht werden, was wiederum zu Lasten der Effizienz ginge.

Nachnamen erweisen sich in dieser Auswertung als trennschärfer (siehe Abbildung 2). Auch die Verzerrung zu Lasten deutscher Namen ist weniger ausgeprägt. Es besteht allerdings, wie oben erwähnt, die Gefahr, dass dies durch Stichprobenfehler zustande kommt. Örtliche Ballungen einzelner Clans führen zu einem hohen Trennschärfemaß bei geringer absoluter Ausbeute, weil die Namen nur in Bielefeld vorkommen. Für überregionale Stichproben sind diese Namen nutzlos. Die Bilanz verschlechtert sich, wie oben gezeigt, bei Verwendung häufiger Familiennamen. Daher soll auf den Einsatz der Nachnamen nicht weiter eingegangen werden.

3.3 Unterscheiden sich Aussiedler mit fremden Namen von denen mit deutschen Namen?

Solange nicht belegt ist, dass sich die Träger deutscher und fremder Namen unterscheiden, spricht nichts gegen die Nutzung überwiegend russischer und polnischer Namen. Man könnte in diesem Fall auf das aufwendige und voraussetzungsvolle Geburtsortsverfahren verzichten und mit einer namensorientierten Auswahl z. B. aus dem Telefonbuch auch zum Ziel kommen, was unter forschungsökonomischen Gesichtspunkten nicht ganz uninteressant wäre. Mit Ergebnissen der Befragung auf Grundlage obiger Stichprobe sollen dazu Anhaltspunkte gesammelt werden.

Betrachtet wird der Vorname. Neben Differenzen im Antwortverhalten kann ein Effekt durch selektives Unit-Non-Response zustande kommen. Für 62 Adressen liegen Gründe der Nicht-Teilnahme vor. Der Anteil der Verweigerer ist unter den Personen mit fremdem höher (62%) als mit deutschen Namen (46%), der Zusammenhang ist aber bei der gegebenen Fallzahl nicht signifikant. Die Betrachtung konzentriert sich daher auf einen Vergleich zwischen 47 Fällen mit deutschen und 35 Fällen mit fremden Namen. Der auffälligste Unterschied ist im durchschnittlichen Einreisezeitpunkt gegeben, der bei deutschem Namen im März 1991 und damit fast zwei Jahre früher als bei fremdem Namen (Februar 1993) liegt. (Das Lebensalter unterscheidet sich wegen der gesetzten Stichprobenkriterien nicht.) Dieser Effekt entspricht der bekannten Tatsache, dass in den späteren Einreisekohorten der Anteil russisch-deutscher Mischehen stark angestiegen ist, während in der frühen Zuwanderungsphase eher Menschen kamen, die ihre ethnische Spezifität als Deutsche bewahrt hatten. In den wichtigsten sozialstrukturellen Dimensionen (Bildung, Einkommen, Erwerbsstatus, Sprachfähigkeit) ließen sich keine signifikanten Unterschiede ermitteln, obwohl für Träger deutscher Namen erhöhte Anteile Vollzeitbeschäftigter und niedrigere Arbeitslosenanteile berechnet wurden. Signifikante Unterschiede (t-Test, $\alpha=5\%$) stellen sich hingegen in Einstellungen und Netzwerkaktivitäten heraus. Wer einen deutschen Vornamen trägt, hat eher das Gefühl, mit autochthonen Deutschen viele Gemeinsamkeiten zu besitzen. Er hat mehr deutsche Vereinskameraden und trifft Freunde aus dem Herkunftsland häufiger (übrigens ebenfalls autochthone deutsche Freunde, aber nicht statistisch signifikant). In der Diskriminierungserfahrung ist

kein signifikanter Unterschied zu entdecken. Resümierend läßt sich festhalten, dass abgesehen von der längeren Aufenthaltsdauer in dieser Auswertung nur punktuelle Differenzen in Einstellungsangelegenheiten sichtbar werden und insgesamt noch keine Strukturabweichungen und damit keine fundamentalen Gründe gefunden wurden, die eine exakte Parität der Namensverteilung zur Repräsentativität einer Stichprobe erforderlich machen. Dies gilt natürlich unter dem Vorbehalt der begrenzten Tragfähigkeit der hier benutzten Stichprobe und muss unter Umständen zukünftig revidiert werden.

4 Vorläufige Bewertung und Ausblick

Die beschriebenen Verfahren zur Identifikation von Aussiedlern aufgrund von Staatsangehörigkeiten und Geburtsort und ggf. Namen sind zwar aufwendig, weil sie historisch-geographische Recherchen voraussetzen und ein gutes Maß nicht-trivialer Datenbankprogrammierung erfordern und weil in den beteiligten Kommunen Datenschutzerfordernisse in höherem Maß als bei üblichen Stichprobenziehungen geprüft werden müssen, wobei zudem beträchtliche Kosten anfallen. Auf lokaler Ebene aber handelt es sich um eine vergleichsweise sehr präzise Methode, die eine Ausschöpfung der Zielgruppe von 95% bei sehr geringem Beifang erwarten lässt.

Entworfen wurde diese Technik zur Stichprobenziehung. Zwei weitere Anwendungsfelder sind aber denkbar:

1. Auch in anderen Datenbeständen, etwa im System der Sozialversicherungen, werden Geburtsorte registriert, wenn auch nicht immer in Verbindung mit mehreren Staatsangehörigkeiten. Die Selektion nach Geburtsort dürfte einen tieferen Einblick in die strukturellen Verteilungen und die Risiken, denen Aussiedler in der Bundesrepublik ausgesetzt sind, bieten, als mit der bisherigen Statistik möglich war.
2. Kommunale Statistiken, etwa zur residentiellen Segregation von Zuwanderern, leiden bekanntermaßen unter den Beschränkungen des Merkmals Staatsangehörigkeit. Das toponomastische Verfahren vermag hier für deutschstämmige Aussiedler Abhilfe zu schaffen.

Stellt die Toponomastik eine Methode *comme il faut* dar, so wird die Forschung oft größere Flächeneinheiten abdecken und deshalb alleine aus befragungsökonomischen Gründen auf die Onomastik zurückgreifen müssen. Dabei sind zwischen Trennschärfe, Ausschöpfung und Repräsentativität Kompromisse einzugehen, mit denen man aber nach derzeitigem Erkenntnisstand wohl leben kann. Es müsste möglich sein, auch mit Listen typischer Aussiedlernamen vor allem aus dem Telefonbuch passable überregionale Stichproben zu generieren.

Eine besondere Herausforderung stellt der Nachwuchs der Aussiedler im juristischen Sinn dar, der für die Sozialwissenschaften nicht weniger relevant sein darf als die im Ausland Geborenen, weil etwa Sprachdefizite im elterlichen Haushalt nicht von heute auf morgen abgelegt werden. Das Geburtsort-Prinzip ist definitionsgemäß für die Stichprobenziehung der zweiten Generation ungeeignet, während Doppelstaatigkeit offensichtlich in gewissem Umfang weiterhin vorliegt. Da sich vermutlich auch Präferenzen für Vornamen rasch ändern, ruhen die Hoffnungen vor allem

darauf, noch eine Weile typische Nachnamen identifizieren zu können. Auch diese werden durch Eheschließungen zunehmend an Trennschärfe verlieren. Vielleicht ist aber der *race-relations-cycle* irgendwann in der Zukunft, wenn auch dieses letzte Kriterium unbrauchbar geworden ist, soweit fortgeschritten, dass sich niemand mehr für Aussiedlernachkommen interessieren muss.

Literatur

- Babka von Gostomski, Ch., 2001: Adressenscreening über Schulen: Erfahrungen beim IKG-Jugendpanel. ZA-Information 49: 29-48.
- Bbeauftragte der Bundesregierung für Migration, Flüchtlinge und Integration, 2005: 6. Bericht über die Lage der Ausländerinnen und Ausländer in der Bundesrepublik Deutschland. http://www.bundesregierung.de/Content/DE/Publikation/IB/Anlagen/ausl_C3_A4nderbericht-6-teil-III,property=publicationFile.pdf (27.6.2007).
- Engels, D., 2000: Aussiedlerinnen und Aussiedler in der Sozialhilfe. Köln: Institut für Sozialforschung und Gesellschaftspolitik (ISG).
- Faulenbach, B., 2002: Die Vertreibung der Deutschen aus den Gebieten jenseits von Oder und Neiße: Zur wissenschaftlichen und öffentlichen Diskussion in Deutschland. Aus Politik und Zeitgeschichte B 51-52: 44-54.
- Gabler, S., 1992: Schneeballverfahren und verwandte Stichprobendesigns. ZUMA-Nachrichten 31: 47-69.
- Haarmann, A., E. Scholz, M. Wasmer, M. Blohm und J. Harkness, 2006: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2004. Mannheim: ZUMA.
- Haisken-DeNew, J. und J. Frick, (Hg.), 2005: Desktop Companion to the German Socio-Economic Panel (SOEP): Version 8.0. Berlin: Deutsches Institut für Wirtschaftsforschung.
- Haug, S. (unter Mitarb. v. P. Schimany), 2005: Jüdische Zuwanderer in Deutschland. Nürnberg: Bundesamt für Migration und Flüchtlinge. Working Papers 3/2005.
- Haug, S. und L. Sauer, 2007: Zuwanderung und Arbeitsmarktintegration von Aussiedlern und Spätaussiedlern. Ergebnisse einer Befragung zur Evaluation des Wohnortzuweisungsgesetzes. Vortragsmanuskript für den Workshop „Integrationschancen von Spätaussiedlern“, Donnerstag, 29.3. – Freitag, 30.3.2007, in der Bundesagentur für Arbeit, Institut für Arbeitsmarkt- und Berufsforschung. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Humpert, A. und K. Schneiderheinze, 2000: Stichprobenziehung für telefonische Zuwandererumfragen – Einsatzmöglichkeiten der Namenforschung (Onomastik). ZUMA-Nachrichten 47: 36-63.
- Humpert, A. und K. Schneiderheinze, 2002: Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahlgrundlage. S. 187-208 in: S. Gabler und S. Häder (Hg.): Telefonstichproben – Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Humpert, A., 2004: Erfahrungen mit Personennamen zur Bildung von Stichproben für Betriebsbefragungen. ZUMA-Nachrichten 54: 141-153.
- Huschka, D., J. Gerhards und G. Wagner, 2005: Messung und Analyse des sozialen Wandel anhand der Vergabe von Vornamen: Aufbereitung und Auswertung des SOEP. Dokumentation der Datenbasis und der Verocoding. Berlin: Institut für Soziologie der Freien Universität Berlin. http://userpage.fu-berlin.de/~gerhards/dokumentefrei/projektdoku_vornamen.pdf (3.7.2007).
- Mammey, U. und R. Schiener, 1996: Das bib-Aussiedlerpanel – methodische Überlegungen zur Repräsentativität. Zeitschrift für Bevölkerungswissenschaft 21: 145-171.
- Mammey, U. und R. Schiener, 1998: Zur Eingliederung der Aussiedler in die Gesellschaft der Bundesrepublik Deutschland: Ergebnisse einer Panelstudie des Bundesinstituts für Bevölkerungsforschung. Opladen: Leske + Budrich (Schriftenreihe des Bundesinstituts für Bevölkerungsforschung 25).
- Salentin, K., 1999: Die Stichprobenziehung bei Zuwandererbefragungen. ZUMA-Nachrichten 45: 115-135.
- Seifert, W., 2001: Gutachten für die „Unabhängige Kommission Zuwanderung“ beim Bundesministerium des Inneren zum Thema: Berufliche Integration von Zuwanderern in Deutschland. Berlin: Bundesministerium des Inneren.

- Schupp, J. und G. Wagner, 1995: Die Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP). Vierteljahrshefte zur Wirtschaftsforschung 64: 16–25.
- Strobl, R. und W. Kühnel, 2000: Dazugehörig und ausgegrenzt. Analysen zu Integrationschancen junger Aussiedler. Weinheim/München: Juventa.
- Swiaczny, F., 2000: Innerstädtische Migration von Aussiedlern: Räumliches Verhalten und Netzwerke als Ursache für Konfliktpotentiale am Beispiel der Stadt Mannheim. Sozialwissenschaften und Berufspraxis 23: 61–85.
- Wetzels, P., D. Enzmann, E. Mecklenburg und Ch. Pfeiffer, 2001: Jugend und Gewalt: Eine repräsentative Dunkelfeldanalyse in München und acht anderen deutschen Städten. Baden-Baden: Nomos.

Korrespondenzadresse: Kurt Salentin
Institut für interdisziplinäre Konflikt- und Gewaltforschung
Universität Bielefeld
33501 Bielefeld
kurt.salentin@uni-bielefeld.de